Capacity of the random Code-Division-Multiple-Access channel with binary inputs

Nicolas Macris, EPFL

Joint work with Satish Korada (Stanford).

References:

- Proc ISIT (2006)
- Proc 45-th Allerton Conf Communication, Control, Computing (2007)
- arXiv:0803.1454 (cs.IT); submitted to IEEE Trans Inf Theory

PLAN

- Code-Division-Multiple-Access setting
- Results in literature for gaussian and binary inputs
- New contributions for binary inputs
- Statistical mechanics formulation
- Interpolation method

DEFINITION OF CDMA

Users $k = 1, ..., K$ transmit $x_k(1)x_k(2)x_k(3)...$ over a common gaussian channel to a single receiver $y(1)y(2)y(3)....$

Users have N "degrees of freedom" available (time slots, frequencies, ...) $j = 1, ..., N$.

- Code division: user k "spreads" its current symbol x_k over the N "degrees of freedom" and transmits the vector

$$\frac{1}{\sqrt{N}} s_{kj} \, x_k, \qquad \mathbb{E}[x_k^2] = 1, \;\; \frac{1}{N} \sum_{j=1}^{N} s_{kj}^2 = 1$$

- Received signal: vector $j = 1, ..., N$

$$y_j = \frac{1}{\sqrt{N}} \sum_{k=1}^{K} s_{kj} \, x_k + \sigma n_j, \qquad \mathbb{E}[n_j^2] = 1$$

- Conventional CDMA: *(Verdu 1986)* spreading sequences are fixed.

$$C = \max_{\prod p_{X_k}} \frac{1}{K} I(\underline{X}; \underline{Y}) = \frac{1}{2K} \log\det(I_K + \sigma^{-2} SS^t)$$

The max is attained at standard gaussian distribution for inputs.

- Random model:*(Verdu-Shamai 1999)* spreading sequence i.i.d standard gaussian s_{kj}; look at $\lim_{K\to\infty}$ with $\frac{K}{N} = \beta$ fixed.

$$C = \underbrace{\frac{1}{2K} \mathbb{E}_{\mathbf{s}} \log\det(I_K + \sigma^{-2} SS^t)}_{\text{can be calculated by RMT}}$$

- Discrete input alphabets: $p_{X_k} = p_k\delta_{-1} + (1 - p_k)\delta_{+1}$

$$\max_{\prod p_{X_k}} \frac{1}{K} I(\underline{X}; \underline{Y})$$

is not known.

- Random model with discrete inputs: i.i.d standard gaussian s_{kj}; look at $\lim_{K\to\infty}$ with $\frac{K}{N} = \beta$ fixed.

$$C = \underbrace{\mathbb{E}_{\mathbf{s}} \frac{1}{K} I(\underline{X}; \underline{Y})}_{\text{no logdet RMT formula}} \quad , \quad p_k = \frac{1}{2}$$

TANAKA'S FORMULA (2001)

Using the formal "replica method" of statistical mechanics one reduces the problem to a variational problem,

$$\lim_{K\to\infty} C = \min_{m\in[0,1]} c(m), \qquad \beta = \frac{K}{N} \ \textit{fixed}$$

For binary input symbols $x_k \in \{+1, -1\}$ and any symmetric distr for s_{kj} with finite second and fourth moments:

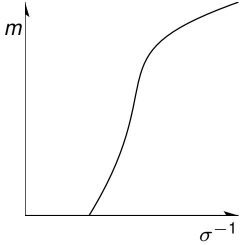
$$c(m) = \frac{\lambda}{2}(1+m) - \frac{1}{2\beta}\ln \lambda\sigma^2 - \int dz\, \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \ln(2\cosh(\sqrt{\lambda}z + \lambda))$$

with

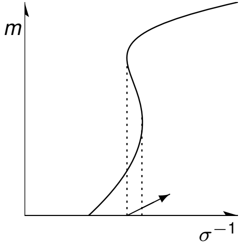
$$\lambda = \frac{1}{\sigma^2 + \beta(1-m)}$$

The minimizer m_* is one of the solutions of

$$m = \int dz \, \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \tanh(\sqrt{\lambda} z + \lambda)), \qquad \lambda = \frac{1}{\sigma^2 + \beta(1-m)}$$



$(\beta < \beta_u)$



$(\beta > \beta_u)$

- $\beta < \beta_u$ unique solution: $m_*(\sigma)$ continuous.
- $\beta > \beta_u$ many solutions: $m_*(\sigma)$ first order phase transition.

The replica method is very powerful...

- "Any" type of input symbol: discrete or continuous. For gaussian inputs,

$$c(m) = \frac{1}{2}\ln(1+\lambda) - \frac{1}{2\beta}\ln\lambda\sigma^2 - \frac{\lambda}{2}(1-m)$$

  $C = \min_{m\in[0,1]} c(m)$ agrees with RMT.

- Unequal powers for users.
- Colored noise
- Communication on CDMA channel with LDPC codes.

(Tanaka, Guo-Verdu, Kabashima-Saad, ...)

RIGOROUS CONTRIBUTIONS (binary inputs)

General assumption: i.i.d

$$p(s_{kj}) = p(-s_{kj}), \qquad p(s_{kj} \geq s) \leq e^{-As^2} \text{ if } s \geq s_0$$

Theorem (S. Korada, N.M 2007)

- *The* $\lim_{K\to\infty} C$ *exists and is equal to* $\lim_{K\to\infty} C_g$ *where* C_g *is the capacity for gaussian* $p(s_{kj})$*.*
- *Tanaka's formula is a lower bound for all* β

$$\lim_{K\to\infty} C \leq \min_{m\in[0,1]} c(m)$$

Montanari and Tse (ITW 2005) sketch the derivation of a lower bound on,

$$\lim_{K\to\infty} \frac{d}{d\sigma} C, \qquad \textit{all } \beta$$

- For $\beta \leq \beta_u$ there is no phase transition and by integrating properly the bound you get $\lim_{K\to\infty} C_K = \min_{m\in[0,1]} c(m)$ for all σ.
- For $\beta \geq \beta_u$ there is a phase transition at $\sigma_c(\beta)$. Above the critical noise their bound is the same than ours. *Well below* the critical noise their bound is the converse one so by combining their result with ours one gets again the equality.

STATISTICAL MECHANICS FORMULATION

To compute $C = \ln 2 - \frac{1}{K}\mathbb{E}_{\mathbf{s}} H(\underline{X} \mid \underline{Y})$ we consider the posterior

$$p(\underline{x} \mid \underline{y}, \mathbf{s}) = \frac{1}{Z(\underline{y}, \mathbf{s})} e^{-\frac{1}{2\sigma^2}\|\underline{y} - N^{-\frac{1}{2}}\mathbf{s}\underline{x}\|^2}$$

with

$$Z(\underline{y}, \mathbf{s}) = \sum_{\underline{x}} e^{-\frac{1}{2\sigma^2}\|\underline{y} - N^{-\frac{1}{2}}\mathbf{s}\underline{x}\|^2}$$

and

$$p(\underline{y} \mid \mathbf{s}) = \sum_{\underline{x}^{input}} \frac{1}{2^K} \frac{e^{-\frac{1}{2\sigma^2}\|\underline{y} - N^{-\frac{1}{2}}\mathbf{s}\underline{x}^{input}\|^2}}{(\sqrt{2\pi}\sigma^2)^N}$$

and

$$p(\mathbf{s}) \qquad i.i.d \quad gaussian$$

This leads to

$$C_K = \ln 2 - \frac{1}{2\beta} - \frac{1}{K}\mathbb{E}_{\underline{Y},\mathbf{s}}[\ln Z(\underline{y},\mathbf{s})]$$

Fundamental object of stat mech "free energy"

$$\frac{1}{K}\ln Z(\underline{y},\mathbf{s})$$

where Z is the "partition function"

$$Z(\underline{y},\mathbf{s}) = \sum_{\underline{x}} e^{-\frac{1}{2\sigma^2}\|\underline{y}-N^{-\frac{1}{2}}\mathbf{s}\underline{x}\|^2} = \sum_{\underline{x}} e^{-\mathcal{H}(\underline{x})}$$

and $\mathcal{H}(\underline{x})$ is the "Hamiltonian" or cost function (log of channel transition probability).

For CDMA the Hamiltonian is

$$\mathcal{H}(\underline{x}) = \frac{\sqrt{\beta}}{2\sigma^2\sqrt{K}} \sum_{k,l}^{K} J_{kl} x_k x_l - \frac{1}{\sigma^2} \sum_{k=1}^{K} h_k x_k + \frac{1}{2\sigma^2} \|\underline{y}\|^2$$

with

$$J_{kl} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} s_{kj} s_{lj}, \qquad h_k = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} y_j s_{kj}$$

- Spins $x_k \in \{-1, +1\}$ are the dynamical degrees of freedom.
- Couplings J_{kl}, h_k are frozen/quenched disorder.

- CDMA is a complicated spin glass model.
- Superficialy similar to the Sherington-Kirkpatrick model

$$\mathcal{H}(\underline{x}) = \frac{1}{\sqrt{K}} \sum_{k,l} J_{kl} x_k x_l \qquad i.i.d \;\; J_{kl} \;\; distr \;\; \mathcal{N}(0, J)$$

- If you change $\mathcal{H}(\underline{x}) \to -\mathcal{H}(\underline{x})$ you get a kind of Hopfield Hamiltonian.
- As for other communications problems: Nishimori gauge symmetry $\to$ replica symmetric solution is expected to be correct.

THE INTERPOLATION METHOD

It was pioneered by Guerra-Toninelli. Based on it Talagrand arrived at a proof of the Parisi formula for SK.

- Takes the replica solution as the favorite guess and tries to find the corresponding "mean field Hamiltonian or channel"
- Construct an interpolating Hamiltonian or channel: $0 \leq t \leq 1$
- Fundamental theorem of calculus

$$\underbrace{\ln Z(1)}_{\text{true system}} = \underbrace{\ln Z(0)}_{\text{mean field syst}} + \int_0^1 dt \, \frac{d}{dt} \underbrace{\ln Z(t)}_{\text{interpolating syst}}$$

- The derivative produces correlation functions with a controllable sign (hopefully).

Step 1. Guessing: $\lim_{K\to\infty} C = \min_{m\in[0,1]} c(m)$

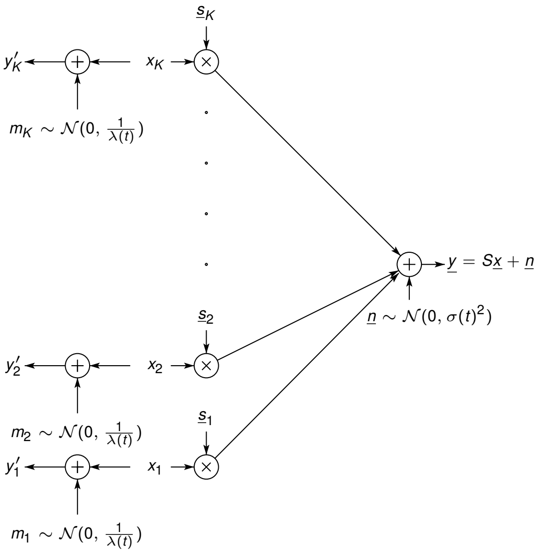
$$c(m) = \frac{\lambda}{2}(1+m) - \frac{1}{2\beta}\ln \lambda\sigma^2 - \underbrace{\int dz\,\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\ln(2\cosh(\sqrt{\lambda}z+\lambda))}_{\textit{almost capacity of BIAWGN}(\lambda^{-1})}$$

Mean field Hamiltonian correspond to K independent BIAWGN channels

$$y'_k = x_k + \lambda^{-1/2} m_k, \qquad m_k \sim \mathcal{N}(0, \lambda^{-1})$$

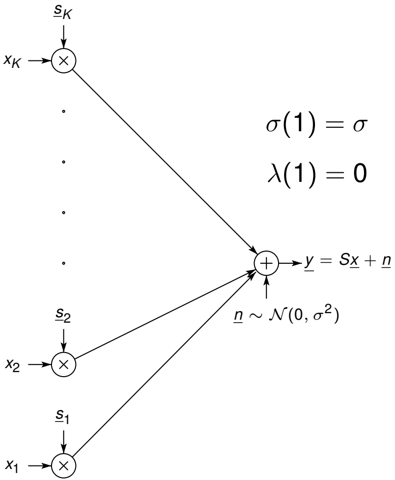
Recall $\lambda^{-1} = \sigma^2 + \beta(1-m)$.

Step 2. Interpolation channel $0 \leq t \leq 1$:



$$\lambda(t) + \frac{1}{\sigma^2(t) + \beta(1-m)} = \frac{1}{\sigma^2 + \beta(1-m)}$$

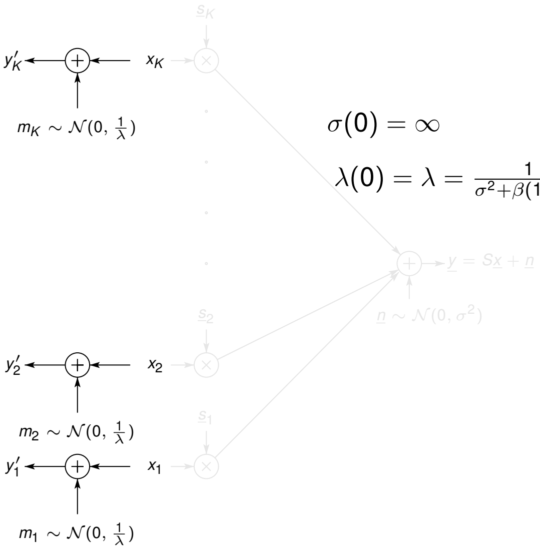
t=1 is the original CDMA channel



$$\sigma(1) = \sigma$$

$$\lambda(1) = 0$$

t=0 are K independent BIAWGN channels



$$\sigma(0) = \infty$$

$$\lambda(0) = \lambda = \frac{1}{\sigma^2 + \beta(1-m)}$$

Capacity of interpolating system

$$\frac{1}{K}\mathbb{E}_{\mathbf{S}} I_t(\underline{X};\underline{Y},\underline{Y}') = \ln 2 - \frac{1}{2\beta} - \mathbb{E}_{\mathbf{S},\underline{Y},\underline{Y}'} \ln Z(t)$$

where

$$Z(t) = \sum_{\underline{x}} e^{-\frac{1}{2\sigma(t)^2}\|\underline{y}-N^{-1/2}S\underline{x}\|^2-\frac{\lambda(t)}{2}\|\underline{y}'-\underline{x}\|^2}$$

and

$$\underline{y} = \frac{1}{\sqrt{N}} S\underline{x}^{input} + \underline{n}, \qquad n_i \sim \mathcal{N}(0, \sigma(t)^2)$$

$$\underline{y}' = \underline{x}^{input} + \underline{m}, \qquad m_k \sim \mathcal{N}(0, \frac{1}{\lambda(t)}).$$

Step 3. Fundamental theorem of calculus:

$$\mathbb{E}_{\mathbf{s},\underline{Y},\underline{Y}'} \ln Z(1) = \mathbb{E}_{\mathbf{s},\underline{Y},\underline{Y}'} \ln Z(0) + \int_0^1 dt \, \frac{d}{dt} \mathbb{E}_{\mathbf{s},\underline{Y},\underline{Y}'} \ln Z(t)$$

This leads to

$$C = c(m) + \int_0^1 dt \, R(t) + o_K(1)$$

with

$$R(t) \leq 0, \quad \textit{for all } 0 \leq t \leq 1$$

Step 4. Sign of remainder: (hard to bring in ratio form)

$$R(t) = -\beta \frac{\sigma'(t)}{\sigma(t)} \frac{\left(\mathbb{E}_{\underline{Y},\underline{Y}',\mathbf{s}}\langle \mu - m\rangle_t\right)^2}{\left(\sigma(t)^2 + \beta(1-m)\right)\left(\sigma(t)^2 + \beta\mathbb{E}_{\underline{Y},\underline{Y}',\mathbf{s}}\langle 1-\mu\rangle_t\right)}$$

where

$$\mu = \frac{1}{K}\sum_{k=1}^{K} x_k^{input} x_k \qquad (magnetization)$$

$$\langle - \rangle_t \ \ is \ \ p_t(\underline{x} \mid \underline{y}, \underline{y}', \mathbf{s}) = \frac{e^{-\mathcal{H}_t(\underline{x})}}{Z_t} \qquad (interpolating\ \ Gibbs)$$

Main ingredients for calculating the remainder

$$\frac{d}{dt}\mathbb{E}_{\underline{Y},\underline{Y}',\mathbf{s}} \ln Z(t) = \mathbb{E}_{\underline{Y},\underline{Y}',\mathbf{s}}\langle \mathit{horrible\,polynomial}(\underline{x},\underline{y},\underline{y}',\mathbf{s})\rangle_t$$

- Integration by parts formula for Gaussian r.v (Guerra)

$$\mathbb{E}[u\varphi(u)] = \mathbb{E}[\varphi'(u)]$$

- Gauge symmetry
- Concentration theorems

Gauge symmetry implies remarkable identities (Nishimori)

$$\mu = \frac{1}{K}\sum_{k=1}^{K} x_k^{input} x_k \qquad \textit{and} \qquad q = \frac{1}{K}\sum_{k=1}^{K} x_k^{(1)} x_k^{(2)}$$

have the same distribution

$$\mathbb{E}_{\underline{Y},\underline{Y}',\mathbf{s}}\langle\, \mu^p \,\rangle_t = \mathbb{E}_{\underline{Y},\underline{Y}',\mathbf{s}}\langle\, q^p \,\rangle_t$$

- Such identities arise in a natural way for various channel models.

Concentration: we need $\mathbb{E}\langle(\mu - \mathbb{E}\langle\mu\rangle)^2\rangle \to 0$

Applying concentration thms for Lipschiz functions of i.i.d gaussians:

Theorem

For $p(s_{kj})$ standard gaussian:

$$\mathbb{P}[|\ln Z(\underline{y}, \mathbf{s}) - \mathbb{E}_{\underline{Y},\mathbf{s}} \ln Z(\underline{y}, \mathbf{s})| \geq \epsilon K] \leq e^{-\alpha(\beta,\sigma)\epsilon^2\sqrt{K}}$$

$$\mathbb{P}[|I(\underline{X}; \underline{Y}) - \mathbb{E}_{\mathbf{s}} I(\underline{X}; \underline{Y})| \geq \epsilon K] \leq e^{-\alpha(\beta,\sigma)\epsilon^2 K}$$

Remark that

$$\langle \mu \rangle = \frac{\partial}{\partial u} \ln \underbrace{\sum_{\underline{x}} e^{-\mathcal{H}(\underline{x}) + u \sum_k x_k^{input} x_k}}_{Z_u}$$

Turn concentration of $\ln Z_u$ into

Corollary

Fix $\epsilon > 0$. For Lebesgue almost every $u > \epsilon$,

$$\lim_{K \to \infty} \int_0^1 dt \mathbb{E} \langle (\mu - \mathbb{E}\langle \mu \rangle_{t,u})^2 \rangle_{t,u} = 0$$

In reality the u- perturbation is more subtle because our methods do not afford to break the channel symmetry.

CONCLUSION

- Other sorts of interpolations allow to prove that lim C_K exists and is independent of distribution of spreading sequence.
- More general input distributions, Unequal powers of users, MIMO , CDMA with users using LDPC Codes, Coloured noise.
- Main open question: lower bound on Capacity. We can do this for simpler special spin glasses with gauge symmetry.
- For LDPC codes over BMS channels the interpolation method has been developed but the lower bound is also missing.